Generating digital twins of mesoporous silica by graph-based stochastic microstructure modeling

Benedikt Prifling\textsuperscript{a,*}, Matthias Neumann\textsuperscript{a}, Dzmitry Hlushkou\textsuperscript{b}, Christian Kübel\textsuperscript{c}, Ulrich Tallarek\textsuperscript{b}, Volker Schmidt\textsuperscript{a}

\textsuperscript{a}Institute of Stochastics, Ulm University, 89069 Ulm, Germany
\textsuperscript{b}Department of Chemistry, Philipps-Universität Marburg, 35032 Marburg, Germany
\textsuperscript{c}Institute of Nanotechnology & Karlsruhe Nano Micro Facility, Karlsruhe Institute of Technology, 76344 Eggenstein-Leopoldshafen, Germany

Abstract

Silica monoliths are hierarchically structured, versatile materials that are widely used in analytical separation science, e.g., liquid chromatography. Their functionality strongly depends on the 3D morphology of their macropore and mesopore spaces. In the present paper, we consider three differently manufactured silica monolith samples, where the process conditions of their hydrothermal treatment (affecting, e.g., mesopore size) have been varied, and present a parametric stochastic 3D microstructure model that is able to generate digital twins of the resulting mesopore spaces. The model, which is based on random point processes and relative neighborhood graphs, theoretically guarantees the complete connectivity of both, the silica phase and the mesopores. The parametric model is fitted to electron tomographic image data. For this purpose, we optimize a cost function that is based on empirically derived relationships between model parameters and volume fraction, mean geodesic tortuosity and constrictivity. Validation is performed regarding further microstructure characteristics, which are not used for model fitting, and regarding effective diffusivity, which is numerically simulated by a particle-tracking algorithm based on random walks.

Keywords: Stochastic microstructure modeling, amorphous silica, random point process, relative neighborhood graph, effective diffusivity, random walk

1. Introduction

Silica monoliths combine low bulk density, high porosity, and mechanical stability with a large surface area. Thus they are versatile materials being widely used in diverse industrial applications for chemical separations, filtering, heterogeneous catalysis, gas adsorption, etc. [1, 2, 3, 4, 5, 6]. Silica monoliths consist of the silica skeleton and pores, where the latter can be subdivided into interskeleton macro- and intraskeleton mesopores. The original process to prepare macro-mesoporous silica monoliths was introduced in [7]. It consists of sol-gel transition with concurrent, polymer-induced phase separation yielding wet, macro-microporous gels, followed by hydrothermal treatment of the gel to widen the micropores to mesopores through surface etching. Silica monoliths

\*Corresponding author. Email: benedikt.prifling@uni-ulm.de. Phone: +49 731 50 - 23555. Fax: +49 731 50 - 23649.
with different properties of the pore space can be synthesized by varying the content of the silica precursor, pH of the starting gel, chemical structure, molecular weight, and content of the polymer, as well as the conditions of the hydrothermal treatment.

While the highly permeable macropores allow for efficient advective transport by pressure-driven flow, the mesopores enable fast transport of solutes to and from active sites inside the skeleton \[2, 5, 8\]. Moreover, the mesoporous skeleton provides a large surface for active sites accessible by diffusion. However, when the interfacial kinetics (sorption, reaction) are fast, hindered diffusion through the network of mesopores inside the monolith skeleton can become the limiting factor to the process performance and efficiency. Thus, a better understanding of relationships between the morphology of mesopores and the corresponding effective properties is crucial to optimize processes controlled by hindered diffusion. Due to this strong impact of the microstructure on effective diffusivity \[9, 10\], identifying preferable morphologies has the potential to significantly improve functional properties of silica monoliths.

For this purpose, the 3D morphology of silica monoliths can be investigated by means of tomo-
graphic imaging and a subsequent statistical analysis of the 3D microstructure. Since manufac-
turing and the subsequent imaging process for a large number of samples is expensive in time and costs, parametric stochastic microstructure modeling is a powerful tool to overcome these limi-
tations via an approach called virtual materials testing. More precisely, a parametric stochastic 3D microstructure model is fitted to tomographic image data, which allows for the generation of digital twins, \textit{i.e.}, virtual microstructures with similar structural properties compared to those observed by tomographic imaging. Note that in materials research the segmented tomographic reconstruction is frequently considered as a digital twin of the 3D microstructure, whereas in the present paper the notion of digital twins refers to realizations of the stochastic 3D microstructure model. By systematic variation of the model parameters, a large range of virtual microstructures can be generated on the computer which are still realistic, but differ from the samples used for imaging. These virtual microstructures serve as input for numerical simulations of functional properties, which makes it possible to empirically derive microstructure-property relationships. Virtual materials testing has been applied, \textit{e.g.}, to quantify the influence of microstructure on effective conductivity and permeability in \[11, 12, 13, 14, 15\].

In the present paper, we perform the first step for virtual materials testing of amorphous, meso-
porous silica from hierarchical, macro-mesoporous silica monoliths. This means that we present a graph-based stochastic microstructure model which enables us to generate digital twins of the mesopore space exhibiting different morphologies. For this purpose, we introduce a new parametric stochastic microstructure model, which is a modification of the model used in \[16\] for applications in solid oxide fuel cells. The model is fitted to tomographic image data of three differently processed silica monolith samples described in \[17\]. These monoliths stem from the same sol-gel processing step, but were subjected to varied conditions during the postgelational, hydrothermal treatment step, resulting in different mesopore size, mesoporosity, and surface area. In the following, the three samples are denoted as Si12, Si21 and Si26. The number corresponds to the mean meso-
pore diameter (in nanometer) based on nitrogen physisorption analysis \[17\]. Tomographic image data have been obtained by scanning transmission electron microscopy (STEM), see \[18, 19\] for an overview. The voxel sizes as well as the overall sizes of the 3D images are listed in Table 1.

The remainder of this paper is organized as follows. In Section 2, we describe several quantities that
are used for characterizing the 3D morphology of the mesoporous silica. The parametric stochastic microstructure model is presented in Section 3. Model fitting to tomographic image data for the generation of digital twins is described in detail in Section 4. Finally, the fitted microstructure model is validated with respect to structural and functional properties, see Section 5. We conclude this paper with a summary of the main results in Section 6.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Si12</th>
<th>Si21</th>
<th>Si26</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size in x-direction [nm]</td>
<td>227.5</td>
<td>224</td>
<td>270</td>
</tr>
<tr>
<td>Size in y-direction [nm]</td>
<td>227.5</td>
<td>224</td>
<td>270</td>
</tr>
<tr>
<td>Size in z-direction [nm]</td>
<td>175</td>
<td>175</td>
<td>135</td>
</tr>
<tr>
<td>Voxel size [nm]</td>
<td>0.35</td>
<td>0.5</td>
<td>0.5</td>
</tr>
</tbody>
</table>

2. Transport-relevant microstructure characteristics

In this section, we briefly explain the mathematical definitions of several transport-relevant microstructure characteristics which are used in the present paper for model fitting and validation. Note that we consider both the solid and the pore phase as stationary random closed sets, which turned out to be a powerful framework for microstructure characterization based on image data, see, e.g., [20, 21]. The random closed set representing the solid phase is denoted by $\Xi_S$, while $\Xi_P$ denotes the random closed set representing the pore space.$^1$ In the following, we use the subscripts $S$ and $P$ in order to indicate whether the considered microstructure characteristic refers to the solid phase or the pore space.

2.1. Volume fraction and porosity

Obviously, one of the most important characteristics is the volume fraction of the solid phase, denoted by $\varepsilon_S$. More precisely, it holds $\varepsilon_S = \mathbb{E}(\nu_3(\Xi_S \cap [0,1]^3))$, where $\nu_3$ denotes the three-dimensional Lebesgue measure. This quantity is estimated from 3D image data using the point-count method, which has been applied since the 19th century and is, e.g., described in [20]. From the volume fraction of the solid phase, we directly obtain the porosity $\varepsilon_P = 1 - \varepsilon_S$.

2.2. Mean geodesic tortuosity

Besides volume fraction, the so-called mean geodesic tortuosity turned out to be of great relevance when considering transport phenomena in dispersed porous microstructures [22]. This quantity characterizes the length of shortest transportation paths. Note that there exist several definitions of tortuosity in the literature. For a comprehensive overview, the reader is referred to [23, 24, 25].

---

$^1$For technical reasons, the interface between solid and pores is considered to belong to both $\Xi_S$ and $\Xi_P$, i.e., we have $\Xi_P = \text{cl}(\Xi_S^C)$, where $A^C$ and $\text{cl}(A)$ denote the complement and the topological closure of a subset $A \subset \mathbb{R}^3$ of the three-dimensional Euclidean space, respectively.
Further descriptors related to the concept of geometrical path lengths with a special focus on their relationship to transport properties are discussed in the literature \[14, 26, 27\]. Mean geodesic tortuosity of a stationary random closed set \(\Xi\) is a direction-dependent quantity. First, a starting plane and a target plane have to be chosen. Then, mean geodesic tortuosity of \(\Xi\) is defined as the ratio of the expected length of the shortest path from an arbitrary point of the starting plane going through \(\Xi\) to the target plane over the distance between starting and target plane. To estimate mean geodesic tortuosity from image data, we compute shortest paths on the voxel grid using Dijkstra’s algorithm, see \[28\]. For a formal definition of this characteristic in the framework of stationary random closed sets, we refer to \[15\].

2.3. Constrictivity

When considering diffusion processes, the occurrence of bottleneck effects can significantly influence the resulting effective diffusivity. Thus, a further important geometrical characteristic is the constrictivity \(\beta \in [0, 1]\), which quantifies the strength of bottlenecks. This microstructure characteristic has been introduced in \[29\] and has proven to play a major role for several functional materials, for example in solid oxide fuel cells \[30, 31\]. The definition of constrictivity is based on the continuous phase size distribution and a geometrically simulated intrusion porosimetry. The continuous phase size distribution \(\text{CPSD} : [0, \infty) \to [0, 1]\) characterizes the morphology of the pre-defined phase, where \(\text{CPSD}(r)\) is given by the maximum volume fraction of the union of (potentially overlapping) balls of radius \(r\), where these balls have to be completely contained in the considered phase. Note that the value \(\text{CPSD}(r)\) coincides with the volume fraction of the considered phase after a so-called morphological opening using a ball with radius \(r\) as structuring element. In particular, \(\text{CPSD}\) coincides with the opening size distribution \[32, 33\]. The radius \(r_{\text{max}}\) is now defined as the largest radius \(r\) for which \(\text{CPSD}(r)\) is greater or equal than half the volume fraction of the considered phase. In contrast to \(\text{CPSD}\), the simulated intrusion porosimetry \(\text{IP} : [0, \infty) \to [0, 1]\) is a direction-dependent characteristic, where the only difference between \(\text{CPSD}(r)\) and \(\text{IP}(r)\) is the fact that the balls of radius \(r\) are no longer allowed to be placed arbitrarily within the considered phase. Instead, they intrude into the considered phase from a predefined starting plane along the specified direction, which accounts for bottleneck effects. Then, analogously to \(r_{\text{max}}\), the radius \(r_{\text{min}}\) is defined as the maximum radius for which \(\text{IP}(r)\) is greater or equal than half the volume fraction of the considered phase. Finally, constrictivity is defined by \(\beta = (r_{\text{min}}/r_{\text{max}})^2 \in [0, 1]\), where lower values of \(\beta\) correspond to more pronounced bottlenecks. Using the concept of morphological opening \[32, 33\] and the Hoshen-Kopelman algorithm \[34\] allows for an estimation of \(\text{CPSD}\) and \(\text{IP}\) and thus for an estimation of constrictivity from image data. We refer to \[15\] for a formal definition of constrictivity and details regarding its estimation.

2.4. Chord length distribution

A further useful microstructure characteristic is the so-called chord length distribution \[32, 33\]. The chord length in direction \(\varphi \in [0, \frac{\pi}{2}] \times [0, 2\pi]\) of a random closed \(\Xi\) is defined as the random length of a line segment in \(\Xi \cap \ell\), where \(\ell\) denotes the line with direction \(\varphi\) going through the origin. The corresponding distribution is said to be the chord length distribution of \(\Xi\) in direction \(\varphi\). We denote the chord length distribution functions in \(x-, y-\) and \(z\)-direction by \(L_x, L_y\) and \(L_z\) and the corresponding mean chord lengths by \(m_x, m_y, m_z\), respectively. In case that \(\Xi\) is isotropic, the chord length distribution functions are independent of the direction. Thus, the chord length
distributions allow us to characterize potential anisotropy effects. Further information regarding chord length distributions can be found in [20, 21, 35].

2.5. Specific surface area
Next, we consider the specific surface area, which is defined as the mean interface area between solid and pores per unit volume. For estimating this characteristic from 3D image data, we compute a weighted sum that arises from considering all local $2 \times 2 \times 2$ voxel configurations [21]. For this purpose, we use the weights proposed in [36].

2.6. Two-point coverage probability function
A further useful quantity for characterizing random closed sets is the so-called two-point coverage probability function [20, 37]. Note that this quantity is also called covariance function in the literature [21, 32, 33]. For a stationary and isotropic random closed set $\Xi$, the probability $\mathbb{P}(o \in \Xi, x \in \Xi)$ depends only on the distance $r = \|x\|$ from $x \in \mathbb{R}^3$ to the origin $o \in \mathbb{R}^3$. In this case, the two-point coverage probability function $C : [0, \infty) \to [0, 1]$ of $\Xi$ is defined by $C(r) = \mathbb{P}(o \in \Xi, x \in \Xi)$ for arbitrary points $x$ with distance $r$ to the origin. This quantity can be estimated from 3D image data by the Fourier-based approach described in [21]. In the present paper, we consider the two-point coverage probability function of the pore space in amorphous, mesoporous silica, denoted by $C_P$ for model validation. Note that $C_P$ uniquely determines $C_S$, the two-point coverage probability function of the solid phase. In particular, we have $C_S(r) = 1 - 2\varepsilon_P + C_P(r)$.

3. Model Description
In the present paper, we modify the graph-based model proposed in [16], where a parametric modeling approach for the three-phase microstructure of solid oxid fuel cell electrodes is described. Doing so, we obtain a new stochastic microstructure model which allows for the generation of digital twins of amorphous silica on the meso-scale. The general idea is to use $k \geq 1$ stochastically independent random point processes, where $k$ is the number of phases of the underlying material. Afterwards, the points of each point process in three-dimensional space are connected according to some deterministic rule leading to spatial graphs $G_1, \ldots, G_k$. Finally, the random closed set $\Xi_i$ modeling the $i$-th phase is given by the set of those points in $\mathbb{R}^3$ which are closer to $G_i$ than to all the other graphs. In the following, the parametric stochastic 3D microstructure model is described in detail, where we will see later that this type of model is able to ensure the complete connectivity of both phases, solid and pores, which is one of the main morphological features of amorphous, mesoporous silica.

3.1. Random point processes
In general, random point processes are a popular tool from stochastic geometry for modeling irregular spatial point patterns [20, 38]. In particular, parametric families of random point processes are frequently used in the context of microstructure modeling, where materials of interest are among others solid oxid fuel cells [16, 39, 40], organic semiconductors [41], mesoporous alumina [42], solar cells [43], and electrodes in lithium-ion batteries [44]. In [16], the vertices of the spatial graphs $G_1, \ldots, G_k$ are modeled by homogeneous Poisson point processes. In the present paper, this modeling step is generalized by using Matérn cluster processes.
instead, which leads to a more flexible model compared to homogeneous Poisson point processes. The Matérn cluster process has three parameters, i.e., the parent intensity $\lambda_1 > 0$, the child intensity $\lambda_2 > 0$ and the cluster radius $r > 0$, where a Poisson point process with intensity $\lambda_1 > 0$ models a point pattern of cluster centers. This is called the parent point process. For each cluster center, a second homogeneous Poisson point process with intensity $\lambda_2 > 0$ is considered within a ball of radius $r$ around the cluster center. These point processes are called child point processes, the points of which form the points of the resulting Matérn cluster process. In particular, a Matérn cluster process is a special case of a so-called Neyman-Scott process [45]. For a more formal definition, we refer to [46].

The simulation of a Matérn cluster process in a bounded sampling window $W \subset \mathbb{R}^3$ is based on the simulation of Poisson point processes as described in [38]. At first, we determine the cluster centers by simulating the parent process in $W \oplus B(x, r)$, where $B(x, r) = \{ y \in \mathbb{R}^3 : \| x - y \| \leq r \}$ denotes the closed ball with radius $r$ around the center $x \in \mathbb{R}^3$ and $\oplus$ denotes the Minkowski addition [33]. More precisely, for two sets $A, B \subset \mathbb{R}^3$, the Minkowski addition of the sets $A$ and $B$ is defined by $A \oplus B = \{ a + b : a \in A, b \in B \}$. For each cluster center, we simulate the corresponding child point process, which is stochastically independent from all other considered point processes. For more details regarding Matérn cluster processes, we refer to [47, 48].

3.2. Relative neighborhood graph

As already mentioned, the second modeling step is to construct a random spatial graph. For this purpose, we make use of the relative neighborhood graph (RNG), which has been introduced in [49]. This graph can be considered as random spatial graph in case that the set of vertices arise from a random point process $X$ in $\mathbb{R}^3$. The relative neighborhood graph generated by $X$, denoted by $RNG_X$, is the random spatial graph $RNG_X = (X, E_X)$, where the (random) vertex set $X$ is given by the random point process itself. The construction of the edge set $E_X$ is based on a so-called critical region $A \subset \mathbb{R}^3$. More precisely, the critical region between two points $x, y \in \mathbb{R}^3$ is given by $A(x, y) = B(x, \| x - y \|) \cap B(y, \| x - y \|)$. Finally, the (random) edge set $E_X$ is given by $E_X = \{(S_i, S_j) \in X \times X : \#A(S_i, S_j) = 2\}$, i.e., there is an edge between two vertices $S_i, S_j \in X$ if the critical region $A(S_i, S_j)$ contains no further points except of $S_i$ and $S_j$. For an efficient numerical computation of RNGs, one can exploit the fact that the relative neighborhood graph of a finite set of points is a subgraph of its Delaunay triangulation [49]. Computing the Delaunay triangulation is carried out with MATLAB [50], which uses the quickhull algorithm [51]. A survey of theoretical results regarding RNGs can be found in [52]. Finally, note that the relative neighborhood graph is completely connected with probability one if the underlying random point process is given by a Matérn cluster process, see [53].

3.3. Random closed set model

At this point we are able to introduce the random closed set model, which will turn out to be capable to generate digital twins of amorphous, mesoporous silica. In this case we just have two phases: the pore space and a solid phase, i.e., $k = 2$.

Thus, the model is based on two stochastically independent Matérn cluster processes $X_S$ and $X_P$ with parent intensities $\lambda_{S,1}, \lambda_{P,1} > 0$, child intensities $\lambda_{S,2}, \lambda_{P,2} > 0$ and cluster radii $r_S, r_P > 0$. 


respectively. Thus, altogether, the model has a six-dimensional parameter vector 

\[ \theta = (\lambda_{S,1}, \lambda_{P,1}, \lambda_{S,2}, \lambda_{P,2}, r_S, r_P) \in (0, \infty)^6. \]

In a next step, we construct both relative neighborhood graphs \( \text{RNG}_{X_S} \) and \( \text{RNG}_{X_P} \). The random closed set \( \Xi_S(\theta) \), which models the solid phase of the mesoporous silica, is given by

\[ \Xi_S(\theta) = \{ x \in \mathbb{R}^3 : d(x, \text{RNG}_{X_S}) \leq d(x, \text{RNG}_{X_P}) \}. \] (1)

In order to formally define the distance \( d(x, G) \) between a point \( x \in \mathbb{R}^3 \) and a random spatial graph \( G \), we consider the graph \( \text{RNG}_{X_S} \) as a random closed set, i.e., \( \Xi(\text{RNG}_{X_S}) = \bigcup_{(S_i, S_j) \in E_{X_S}} S_i S_j \), where \( S_i S_j = \{ z \in \mathbb{R}^3 : z = \lambda S_i + (1 - \lambda)S_j \text{ for some } \lambda \in [0, 1] \} \) denotes the line segment between \( S_i \) and \( S_j \). The distance between a point \( x \in \mathbb{R}^3 \) and the relative neighborhood graph \( \text{RNG}_{X_S} \) can be now defined as \( d(x, \text{RNG}_{X_S}) = \inf_{y \in \text{RNG}_{X_S}} \|x - y\| \), where \( d(x, \text{RNG}_{X_P}) \) is defined analogously.

Note that the complete connectivity of both RNGs directly implies that the random closed sets \( \Xi_S \) and \( \Xi_P \) are completely connected, which motivates the choice of this model since the solid phase as well as the pore space in mesoporous silica are typically completely connected.

### 4. Model fitting

In this section we describe in detail how the model described in Section 3 is fitted to the three data sets Si12, Si21 and Si26, which have been obtained by scanning transmission electron microscopy (STEM), see [17] for details. At first, we have to preprocess the 3D image data since our stochastic 3D microstructure model is isotropic. In a next step, we describe the numerical optimization procedure to obtain the “optimal” model parameters, where we make use of empirically derived relationships between model parameters and three microstructure characteristics (volume fraction, mean geodesic tortuosity and constrictivity). Since these quantities do not depend on the scale, we rescale the model parameters such that the chord length distribution, and thus the scale, is matched. Finally, a post-processing step is applied in order to further improve the visual accordance between model realizations and tomographic image data.

#### 4.1. Pre-processing

The utilized random closed set model is isotropic. However, by analyzing the mean chord length of the silica phase of all three data sets in \( x-, y- \) and \( z- \)direction, one can observe that the tomographic image data is not isotropic, see Table 2, which is caused by the missing wedge during the imaging procedure [54].

In order to obtain image data, for which the microstructure characteristics described in Section 2 do not depend on the chosen direction, we rescale the 3D binary images in \( x- \) and \( y- \)direction using nearest neighbor interpolation, see [55], where the subsequent procedure is analogously applied to all three tomographic image data sets. In order to calculate the scaling factors in \( x- \) and \( y- \)direction, denoted by \( s_x \) and \( s_y \) respectively, we minimize the \( L^2 \)-distance between the chord length distribution function in \( z- \)direction and the chord length distribution function in \( x- \) or \( y- \)direction after rescaling. The resulting scaling factors are given in Table 2.
Table 2: Mean chord length in x-, y-, and z-direction as well as the scaling factors for each sample.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Si12</th>
<th>Si21</th>
<th>Si26</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_{x,S}$</td>
<td>13.0</td>
<td>16.3</td>
<td>17.6</td>
</tr>
<tr>
<td>$m_{y,S}$</td>
<td>10.9</td>
<td>14.0</td>
<td>15.2</td>
</tr>
<tr>
<td>$m_{z,S}$</td>
<td>15.9</td>
<td>19.2</td>
<td>19.0</td>
</tr>
<tr>
<td>$s_x$</td>
<td>0.838</td>
<td>0.862</td>
<td>0.871</td>
</tr>
<tr>
<td>$s_y$</td>
<td>0.7189</td>
<td>0.754</td>
<td>0.794</td>
</tr>
</tbody>
</table>

4.2. Simulation study

To quantitatively investigate the relationships between the six model parameters $\lambda_{S,1}$, $\lambda_{P,1}$, $\lambda_{S,2}$, $\lambda_{P,2}$, $r_S$ and $r_P$, we generated 324 virtual microstructures in the sampling window $W = [0, 400]^3$ by varying the six parameters of the model described in Section 3. For each of these model realizations, the volume fraction $\varepsilon_S$, the mean geodesic tortuosity $\tau_S$ and the constrictivity $\beta_S$ are computed, i.e., all these quantities are computed with respect to the solid silica phase. Afterwards, we fit a linear model for each of these characteristics to predict them as functions of the model parameters.

To begin with, we investigate the relationship between model parameters and volume fraction. Since $\varepsilon_S \in [0, 1]$, we used the link function $f_1(x) = \log(x/(1-x))$ to ensure that the predicted volume fraction $\hat{\varepsilon}_S$ is always between zero and one. More precisely, we consider the linear model

$$f_1(\hat{\varepsilon}_S) = c_1 \log(\lambda_{S,1}) + c_2 \lambda_{S,2} + c_3 r_S + c_4 \log(\lambda_{P,1}) + c_5 \lambda_{P,2} + c_6 r_P.$$  \hspace{1cm} (2)

It turns out that using $\log(\lambda_{S,1})$ and $\log(\lambda_{P,1})$ as predictors leads to a better fit compared to $\lambda_{S,1}$ and $\lambda_{P,1}$. After computation of the regression coefficients $c_1, ..., c_6$, we apply the inverse of the link function, which is given by $f_1^{-1}(x) = \exp(x)/(1 + \exp(x))$, in order to obtain the predicted volume fraction $\hat{\varepsilon}_S \in [0, 1]$ leading to $R^2_\varepsilon = 0.9872$.

To express $\beta_S \in [0, 1]$ in terms of model parameters, we use again the link function $f_1$ for predicting constrictivity, which results in the model

$$f_1(\hat{\beta}_S) = c_1 \log(\lambda_{S,1}) + c_2 \log(\lambda_{S,2}) + c_3 r_S + c_4 \log(\lambda_{P,1}) + c_5 \log(\lambda_{P,2}) + c_6 r_P$$  \hspace{1cm} (3)

with $R^2_\beta = 0.6466$.

Finally, for relating the model parameters to $\tau_S \in [1, \infty)$, we consider the link function $f_2(x) = \log(x-1)$ together with its inverse function $f_2^{-1}(x) = \exp(x) + 1$. The linear model for predicting mean geodesic tortuosity is given by

$$f_2(\hat{\tau}_S) = c_0 + c_1 \lambda_{S,1} + c_2 \log(\lambda_{S,2}) + c_3 r_S + c_4 \lambda_{P,1} + c_5 \log(\lambda_{P,2}) + c_6 r_P,$$  \hspace{1cm} (4)

which leads to $R^2_\tau = 0.8851$.

The summary statistics for the linear regression models given in (2) to (4) can be found in Table 3 and the corresponding prediction error plots are shown in Figure 1. Even though the goodness of fit with regard to tortuosity and constrictivity is improvable, we will see later on that the model realizations are nevertheless similar in terms of morphological characteristics as well as with regard to effective diffusivity. Furthermore, the $R$-squared statistic will be used as weighting
Table 3: Summary statistics for the linear regression models given in (2), (3), and (4). Note that $R^2$ is computed after applying the inverse of the link function.

<table>
<thead>
<tr>
<th>coefficient</th>
<th>$c_0$</th>
<th>$c_1$</th>
<th>$c_2$</th>
<th>$c_3$</th>
<th>$c_4$</th>
<th>$c_5$</th>
<th>$c_6$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>model in (2)</td>
<td>-</td>
<td>0.7201</td>
<td>0.6402</td>
<td>0.2657</td>
<td>-0.7288</td>
<td>-0.6256</td>
<td>-0.2608</td>
<td>0.9872</td>
</tr>
<tr>
<td>model in (3)</td>
<td>-</td>
<td>0.4287</td>
<td>0.3816</td>
<td>0.1638</td>
<td>-0.4590</td>
<td>-0.3371</td>
<td>-0.1357</td>
<td>0.6466</td>
</tr>
<tr>
<td>model in (4)</td>
<td>-2.7035</td>
<td>-30916</td>
<td>-0.5192</td>
<td>-0.2179</td>
<td>32173</td>
<td>0.4817</td>
<td>0.1989</td>
<td>0.8851</td>
</tr>
</tbody>
</table>

Figure 1: Prediction of volume fraction (left), constrictivity (center) and mean geodesic tortuosity (right). All characteristics have been computed on the solid phase. The red line is the line through the origin with slope one.

factor in Section 4.3, which allows us to put more weight on characteristics that can be predicted more accurately. Note that we are also able to predict volume fraction, mean geodesic tortuosity and constrictivity of the pore space with Equations (2), (3), and (4) by simply interchanging $\lambda_{S,1}$ with $\lambda_{P,1}$, $\lambda_{S,2}$ with $\lambda_{P,2}$ and $r_S$ with $r_P$. These empirically derived relationships between the six model parameters and three microstructure descriptors are used for computing “optimal” model parameters, which is described in the next section.

4.3. Optimization procedure

Our goal is to determine a six-dimensional parameter vector $\theta_{opt}$ such that virtual microstructures generated with $\theta_{opt}$ are similar to the tomographic image data. Hence, the cost function $g(\theta)$ is based on five image characteristics, i.e., the quantities $\varepsilon_S$, $\tau_S$, $\tau_P$, $\beta_S$, and $\beta_P$. When using a hat over the symbol of the considered quantity, we refer to the predicted value of this quantity obtained by the regression models of the previous section. For a given parameter vector $\theta = (\lambda_{S,1}, \lambda_{P,1}, \lambda_{S,2}, \lambda_{P,2}, r_S, r_P) \in (0, \infty)^6$, the cost function is defined as

$$g(\theta) = R^2_\varepsilon \left( \frac{\hat{\varepsilon}_S(\theta) - \varepsilon_S}{\varepsilon_S} \right)^2 + R^2_\tau \left( \frac{\hat{\tau}_S(\theta) - \tau_S}{\tau_S} \right)^2 + R^2_\tau \left( \frac{\hat{\tau}_P(\theta) - \tau_P}{\tau_P} \right)^2 + R^2_\beta \left( \frac{\hat{\beta}_S(\theta) - \beta_S}{\beta_S} \right)^2 + R^2_\beta \left( \frac{\hat{\beta}_P(\theta) - \beta_P}{\beta_P} \right)^2.$$  

The cost function $g(\theta)$ is numerically minimized in MATLAB [50] using the interior-point algorithm [56, 57, 58]. For the resulting “optimal” model parameters, the corresponding model realizations reflect the main features of the tomographic image data well with two exceptions. The first problem concerns the scale of the simulated microstructures, which will be addressed in the next section. The second issue refers to the “smoothness” of the simulated microstructures, which will be increased by a post-processing step.
Table 4: Model parameters (rounded to two decimal places) of the fitted microstructure model for each sample.

<table>
<thead>
<tr>
<th></th>
<th>$\lambda_{S,1}[nm^{-3}]$</th>
<th>$\lambda_{P,1}[nm^{-3}]$</th>
<th>$\lambda_{S,2}[nm^{-3}]$</th>
<th>$\lambda_{P,2}[nm^{-3}]$</th>
<th>$r_S[nm]$</th>
<th>$r_P[nm]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si12</td>
<td>$6.55 \cdot 10^{-4}$</td>
<td>$6.55 \cdot 10^{-4}$</td>
<td>$3.64 \cdot 10^{-3}$</td>
<td>$4.36 \cdot 10^{-3}$</td>
<td>3.51</td>
<td>3.57</td>
</tr>
<tr>
<td>Si21</td>
<td>$2.60 \cdot 10^{-4}$</td>
<td>$2.60 \cdot 10^{-4}$</td>
<td>$1.10 \cdot 10^{-3}$</td>
<td>$1.74 \cdot 10^{-3}$</td>
<td>4.62</td>
<td>4.87</td>
</tr>
<tr>
<td>Si26</td>
<td>$2.20 \cdot 10^{-4}$</td>
<td>$1.82 \cdot 10^{-5}$</td>
<td>$4.65 \cdot 10^{-4}$</td>
<td>$1.53 \cdot 10^{-3}$</td>
<td>4.64</td>
<td>5.08</td>
</tr>
</tbody>
</table>

4.4. Rescaling of model parameters

In order to force the virtually generated microstructures to lie on the same length scale as the tomographic image data, we make use of a helpful theoretical property of the proposed model. More precisely, the volume fraction, mean geodesic tortuosity and constrictivity of both phases do not change, when using the parameter vector $\theta = (s^3\lambda_{S,1}, s^3\lambda_{P,1}, s^3\lambda_{S,2}, s^3\lambda_{P,2}, sr, sr_p)$ with an arbitrary scalar $s > 0$ instead of $\theta$. In contrast to this, the mean chord lengths in $x$-, $y$- and $z$-direction depend on $s$, i.e., it holds that $m_x(\Xi_S(\theta)) = \frac{1}{s} \cdot m_x(\Xi_S(\theta))$. Of course, this formula is also true, when computing the mean chord length in $y$- and $z$-direction, respectively. Thus, we are able to rescale the parameter vector $\theta$ in such a way that the mean chord lengths coincide without changing volume fraction, mean geodesic tortuosity and constrictivity. More precisely, the optimal scaling factor $s_{opt}$ for each of the three samples is computed as

$$s_{opt} = \frac{1}{3} \left( \frac{\bar{m}_{x,S}}{m_{x,S}} + \frac{\bar{m}_{y,S}}{m_{y,S}} + \frac{\bar{m}_{z,S}}{m_{z,S}} \right),$$

where the tilde indicates an average of ten model realizations and the symbols without tilde correspond to tomographic image data. The values of $s_{opt}$ are given by 0.97858 (Si12), 1.0269 (Si21) and 0.9843 (Si26), respectively.

4.5. Post-processing

After rescaling the model parameters, we obtain virtual microstructures, which reflect the 3D morphology of tomographic image data better, but still with one exception. More precisely, in the simulated microstructures the interface between the mesopores and the silica phase is not as “smooth” as in the tomographic image data. Thus, we apply a post-processing step as in [12, 40] in order to smooth the simulated binary images. For this purpose, a Gaussian filter is used [55]. To be more precise, we apply a Gaussian filter with a standard deviation of $\sigma_x = \sigma_y = \sigma_z = 4$ to the model realizations and apply a global threshold, where the threshold is chosen in such a way that the volume fraction of the smoothed model realization coincides with the target volume fraction $\varepsilon_S$. Note that this step is carried out on a bounded sampling window. The effect of smoothing is visualized in Figure 2.

A visual comparison between realizations of the fitted stochastic 3D microstructure model and tomographic image data is shown in Figure 3.
Figure 2: Two-dimensional cutout (150 nm × 100 nm) of a model realization of sample Si12 before (left) and after smoothing (center) as well as a cutout of the tomographic image data (right) for comparison.

Figure 3: Volume renderings of tomographic image data (top row) and realizations of the fitted microstructure model (bottom row) for the samples Si12 (left), Si21 (center) and Si26 (right).

5. Model validation

After we have fitted the stochastic 3D model to the samples Si12, Si21 and Si26, respectively, we now validate the model in two different ways. On the one hand, we compare microstructure characteristics of tomographic image data with those of model realizations. On the other hand, we simulate the effective diffusivity $D_{\text{eff}}$ by means of a particle-tracking algorithm based on random walks [59] and are thus able to compare simulated and tomographically measured 3D microstructures with respect to their functionality. Note that all characteristics used within this section are averaged results based on ten model realizations for each sample, where the size of the sampling window as well as the voxel size is given in Table 1.
Table 5: Comparison between tomographic image data and simulated microstructures with regard to various image characteristics. The values in square brackets correspond to the standard deviations estimated from ten model realizations.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Si12 (%)</th>
<th>Si21 (%)</th>
<th>Si26 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tomography:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\varepsilon_S$</td>
<td>42.33</td>
<td>35.08</td>
<td>29.82</td>
</tr>
<tr>
<td>Simulation:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\varepsilon_S$</td>
<td>42.33 (0.05)</td>
<td>35.06 (0.04)</td>
<td>29.83 (0.04)</td>
</tr>
<tr>
<td>Tomography:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specific surface area [nm$^{-1}$]</td>
<td>0.1121</td>
<td>0.0727</td>
<td>0.0581</td>
</tr>
<tr>
<td>Simulation: Specific surface area [nm$^{-1}$]</td>
<td>0.1033 (0.0006)</td>
<td>0.0683 (0.0007)</td>
<td>0.0520 (0.0006)</td>
</tr>
<tr>
<td>Tomography:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau_P$</td>
<td>1.055</td>
<td>1.055</td>
<td>1.043</td>
</tr>
<tr>
<td>Simulation: $\tau_P$</td>
<td>1.059 (0.003)</td>
<td>1.045 (0.002)</td>
<td>1.037 (0.002)</td>
</tr>
<tr>
<td>Tomography:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau_S$</td>
<td>1.110</td>
<td>1.158</td>
<td>1.179</td>
</tr>
<tr>
<td>Simulation: $\tau_S$</td>
<td>1.097 (0.006)</td>
<td>1.151 (0.016)</td>
<td>1.212 (0.0356)</td>
</tr>
<tr>
<td>Tomography:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_P$</td>
<td>0.602</td>
<td>0.676</td>
<td>0.593</td>
</tr>
<tr>
<td>Simulation: $\beta_P$</td>
<td>0.621 (0.020)</td>
<td>0.661 (0.024)</td>
<td>0.637 (0.029)</td>
</tr>
<tr>
<td>Tomography:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_S$</td>
<td>0.600</td>
<td>0.505</td>
<td>0.461</td>
</tr>
<tr>
<td>Simulation: $\beta_S$</td>
<td>0.531 (0.013)</td>
<td>0.433 (0.041)</td>
<td>0.397 (0.039)</td>
</tr>
</tbody>
</table>

5.1. Structural validation

To begin with, we consider the volume fraction of the silica phase. As it can be seen in Table 5, this quantity fits perfectly, which is not surprising since the smoothing is carried out in such a way that the volume fraction $\varepsilon_S$ observed in tomographic image data is matched.

Next, we examine the goodness of fit with regard to specific surface area. Table 5 indicates that specific surface area is slightly underestimated by the model. Note that this quantity is not exactly matched, potentially because the post-processing step lowers the specific surface area. Nevertheless, the relative error with regard to this quantity has significantly decreased by introducing the post-processing step.

Next, we compare simulated and tomographically measured 3D microstructures with respect to mean geodesic tortuosity, which has been defined in Section 2.2. As one can easily infer from Table 5, relative errors between 0.4% and 2.8% indicate that mean geodesic tortuosity in both phases is nicely matched by the model.

The last scalar quantity which we consider in this section is constrictivity. With regard to the pore space, which is the phase where diffusion processes take place, we observe only minor differences between the constrictivity of simulated and tomographically measured 3D microstructures. When considering the silica phase instead, the model realizations possess a significantly lower constrictivity, which might be caused by the fact that constrictivity can not be predicted as accurately as volume fraction and mean geodesic tortuosity.
In addition to the above mentioned scalar quantities, we will now validate the model with regard to function-valued characteristics. To begin with, we consider the continuous size distributions of both phases, which are shown in Figure 4. All in all, the main differences between the 3D morphologies of the three samples are reflected by the model. In addition, the increase of the mean pore diameter from sample Si12 via Si21 to Si26 is reproduced in the model, which corresponds to a shift of the CPSD to the right. However, the simulated microstructures do not exhibit large pores, i.e., the continuous pore size distribution is decreasing faster to zero compared to tomographic image data.

Figure 4: Continuous phase size distribution of the silica phase (left) and the pore space (right) for each sample.

As already mentioned in Section 2, one can take the strength of bottleneck effects into account by simulating intrusion porosimetry. Figure 5 shows that the behavior of this quantity with respect to both phases is nicely captured by the fitted models. In addition, the characteristic bottleneck radius, i.e., the radius for which a significant drop in the IP can be observed, is also matched by the simulated mesoporous silica.

Figure 5: Simulated intrusion porosimetry of the silica phase (left) and the pore space (right) for each sample.
Next, we consider the chord length distribution function of the silica phase, where the corresponding plots are shown in Figure 6. First, one can observe that the chord length distribution functions in \( x \)-, \( y \)- and \( z \)-direction of simulated 3D microstructures as well as of tomographic image data coincide. This is not surprising since the tomographic image data has been appropriately rescaled and the model is isotropic by construction. In addition, the shape of all plots in Figure 6 is very similar to each other even though the mean chord length has been used as the only information for rescaling the model parameters in Section 4.4.

![Figure 6: Chord length distribution functions of the silica phase for the samples Si12 (left), Si21 (center) and Si26 (right).](image)

Finally, we consider the two-point coverage probability function of the mesopores, which has been introduced in Section 2.6. Figure 7 only shows minor deviations between the plots of this characteristic for simulated and tomographically measured 3D microstructures.

![Figure 7: Two-point coverage probability function of the pore space for the samples Si12 (left), Si21 (center) and Si26 (right).](image)

5.2. Functional validation
After having validated the model with regard to various microstructure characteristics, we now validate the model with respect to an important functional property. More precisely, the effective diffusion coefficients \( D_{\text{eff}} \) for the simulated and tomographically reconstructed microstructures were determined with a technique based on random walks and particle tracking (RWPT) [60, 61], which has already been used in previous papers [26, 62, 63]. Briefly speaking, a large number \( N = 5 \cdot 10^6 \) of point-like particles (tracers) was randomly and uniformly distributed in the pore space of the
analyzed 3D microstructure. Then, during each time step $\Delta t$ of the simulation, the displacement of every tracer due to random diffusive motion was calculated from a Gaussian distribution with zero mean and standard deviation $(2D_0 \Delta t)^{1/2}$ along each Cartesian coordinate, where $D_0$ is the diffusion coefficient of tracers in free space, sometimes called intrinsic diffusivity. The random motion was simulated assuming that there are no interactions among the tracers and no passive interactions of the tracers with the pore walls, which was handled through a multiple-rejection boundary condition at the solid surface [64, 65]. When a tracer hits the pore wall during an iteration, the displacement was rejected and recalculated until the target tracer position was in the pore space. In addition, mirror boundary conditions have been implemented to ensure that there are no abrupt changes when a marker “leaves” the observation window. The time step $\Delta t$ of the RWPT-simulation was adjusted such that the mean diffusive displacement did not exceed $\Delta h_{10}$, where $\Delta h$ is the spatial resolution of the voxel grid. The effective diffusion coefficient was evolved from

$$D_{\text{eff}} = \lim_{t \to \infty} \frac{1}{6N} \frac{\partial}{\partial t} \sum_{i=1}^{N} \| r_i(t) - r_i(0) \|^2,$$

where $r_i(t)$ denotes the position of the $i$-th tracer at time $t$. A normalized characteristic derived for effective diffusivity is the so-called $M$-factor, which is defined by $M = D_{\text{eff}} / D_0 \cdot \varepsilon P \in [0, 1]$. As it can be seen in Figure 8, the $M$-factor is slightly overestimated by the proposed 3D microstructure model. This might be caused by the slight underestimation of the specific surface area through the model, i.e., the surface between the pore space and the solid phase is less rugged, which possibly leads to less hindered diffusion processes and thus a slightly larger $M$-factor. Nevertheless, the mean relative error with respect to the $M$-factor equals 2.8% (Si12), 2.7% (Si21) and 2.6% (Si26), respectively, which shows that the digital twins of the considered silica samples generated by the stochastic microstructure model lead to a similar diffusion behavior compared to the tomographic image data. This conclusion is also in good accordance with the prediction formula presented in [30], where it has been shown that the $M$-factor can be reliably predicted by volume fraction, mean geodesic tortuosity and constrictivity via

$$\widehat{M} = \varepsilon P^{1.67 - 0.48 \beta P} \beta P^{5.18}.$$

More precisely, the relative errors between the simulated and the predicted $M$-factor computed on tomographic image data are equal to 3.7% (Si12), 2.3% (Si21) and 5.2% (Si26), respectively. The average of the relative error of ten model realizations equals 4.1% (Si12), 4.5% (Si21) and 5.8% (Si26), respectively. Note that the predicted $M$-factor always slightly underestimates the numerically simulated $M$-factor. For investigations regarding the relationship between microstructure characteristics and the $M$-factor of the interskeleton macropores in silica monoliths, we refer to [66].

6. Conclusion

In the present paper, we considered three differently processed samples of amorphous, mesoporous silica from the (mesoporous) skeleton of hierarchically structured, macro-mesoporous silica monoliths. We have presented a parametric random closed set model, which is able to generate digital
twins of these samples, where both pore space and solid phase were completely connected. The model has been fitted to 3D image data obtained by scanning transmission electron microscopy (STEM). Model fitting has been carried out by taking advantage of empirically derived relationships between the model parameters and three microstructure descriptors, i.e., volume fraction, mean geodesic tortuosity and constrictivity. Moreover, the proposed model has been validated with respect to morphological properties, which have not been used for model fitting, and with regard to effective diffusivity. The very good accordance between tomographic image data and virtually generated morphologies shows that we are really able to generate digital twins, i.e., virtual, but realistic mesoporous silica materials. Thus, by systematic variation of model parameters the proposed model can be used in the future to generate a broad range of further virtual mesoporous silica morphologies, which serve as a geometry input for numerical simulations of effective diffusivity or other functional properties. This approach, called virtual materials testing, allows us to systematically investigate microstructure-property relationships and to detect preferable morphologies with optimized properties.

**Data availability**

The raw/processed data required to reproduce these findings cannot be shared at this time as the data also forms part of an ongoing study.

**References**


